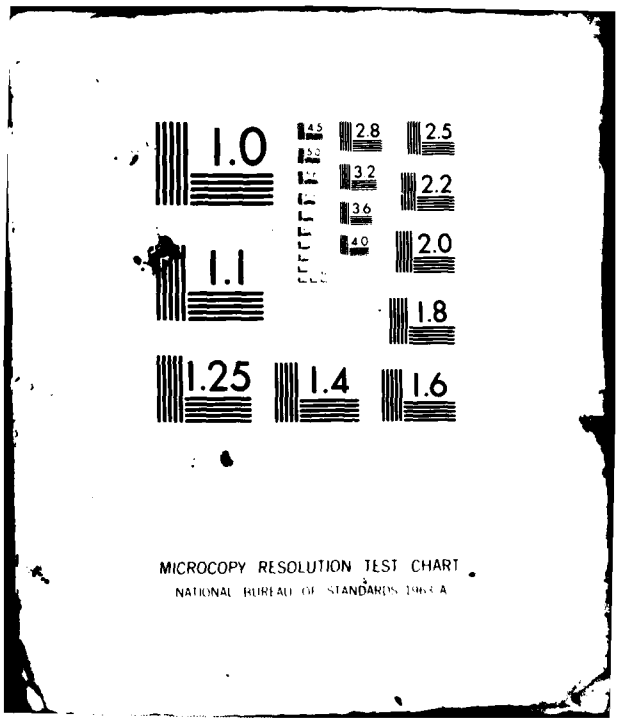
AD-A112 955 NATIONAL RESEARCH COUNCIL WASHINGTON DC NUMERICAL DA--ETC F/G 5/2
REPORT OF AD HOC NDAB SUBGROUP.(U)
MAR 82 NB80NADA1036

UNCLASSIFIED NL

END
DATE FILMED
5-82
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

Numerical Data Advisory Board
Assembly of Mathematical and Physical Sciences
National Academy of Sciences/National Research Council

Report of

Ad Hoc NDAB Subgroup

to the

Defense Technical Information Center
Cameron Station
Alexandria, Virginia

Work performed under Contract No. MDS0MADA1036. This contract includes an interagency transfer from the Defense Technical Information Center, Defense Logistics Agency, Department of Defense, for the performance of this work. Additional support was received from the Army Materials and Mechanics Research Center.

March 1, 1982

Preface

This report is the result of a request made by Hubert Sauter, Administrator, Defense Technical Information Center (DTIC), and is based on 1) a presentation by Sauter and DoD staff at the Numerical Data Advisory Board (NDAB) meeting of 13-14 July 1981, and 2) a follow-up meeting on 31 July at DTIC in response to discussions at the NDAB meeting, attended by DTIC staff and a subgroup consisting of Samuel Etris, Donald King, and Cynthia Carter (NRC staff). The intent of these meetings was to initiate discussions and provide a preliminary transmittal to DTIC, which is seeking NDAB advice in establishing a more numerically oriented program.

Therefore, this report covers only some of the most significant points for DTIC consideration. It identifies a number of broad, long-range data issues as well as specific tasks that require immediate attention. It makes recommendations in those areas of highest priority that should be addressed even in the event of reduced budget. Some of these may result in cost saving. Other subjects identified in this report should be addressed to the extent that a less constrained budget may allow. DTIC must anticipate increased computer sophistication, automated methods in design and manufacturing, improved understanding of artificial intelligence capabilities, and as yet unprecedented pathways of numerical data from laboratory of origin to end user. Development of forefront data/information utilization and transfer methods is prerequisite for development of high defense technology.

Recommendations for immediate action:

The subjects identified in this report that should be given the highest priority at this time are summarized here. The alphanumeric notations in parentheses refer to the sections and paragraphs of the text in which the subject is further addressed.

- The question of improving efficiency of data flow (1; 4):
  - role of data repositories (1b);
  - computerization of numerical data handling (4a, b, c, d);
  - cost/benefit consideration of tagging and flagging or alternate numerical data transfer methods (1g).
- Quality/efficiency of existing Information Analysis Centers (1d; 2a, b).
- Determination of DoD data requirements at present and for anticipated high technology development. Criteria for priority setting should include investment/payoff projections (see 1e, f).
- Completion of the compilation of an index of information activities and sources within DoD. This is considered to become a valuable resource and NDAB encourages this work. The entries must be indexed adequately to assure utility of this compilation (see also 2f).

Overview

The Defense Technical Information Center (DTIC) of the Department of Defense (DoD) currently has four specific tasks:

1. The DoD Technical Report Literature — an indexed bibliographic system with hard copy distribution capability

2. Ongoing Research — an indexed bibliographic database containing summaries for each project

3. Program Planning Database

4. Independent Research and Development — a DoD research funding program

DTIC also administers the Defense Logistics Agency's Information Analysis Centers (IACs) and has been given the broad task of improvement and maintenance of the flow of information for DoD and to consider use of information by R&D managers.

The total DTIC mandate can thus be considered to be rather broad. DoD is an agency with a unique data resource in that it encompasses a microcosm of the data and information with which the entire U.S. deals (e.g. hard sciences, geological, environmental, medical, economic, meteorological, sociological). Thus, DTIC, with this unique resource, is in a position to study a considerable number of data problems that are relevant to the nation as a whole. If DTIC chooses to make it so, the outcome of its own studies could have major impact on information quality and flow on a national scale. For example, problems with interlaboratory data flow among experimentalists or interdisciplinary flow from generator, evaluator, to end user are essentially the same for DoD as they are for the nation as a whole. However, they are critically important for both defense hardware (weapons, military instrumentation) development and strategic decisions and therefore more acute for DoD than for several other U.S. sectors.

The suggested topics and recommendations in this transmittal are only a preliminary summary of some important items for DTIC (no order of importance). At this time, the total range of data issues has not yet been assessed for DoD by NDAB nor has a complete listing been achieved at this time. NDAB is available to assist DTIC in further development of this listing, both in magnitude and depth, and in furthering program changes to establish DTIC's numerical data capabilities.

Topics for Consideration:

1. DTIC role in numerical data:

Several topics can be listed at this time:

a. Determine the rate at which data measured in the laboratory become available in raw and in evaluated form to the user. Alternate shorter routes to data flow should be identified.

b. Determine the need for data repositories. The report literature does not capture all technical data generated in DoD laboratories or by DoD contractors. It would be useful to choose a technical subject area in order to study, for example, an alternative approach of quite brief reports tied in with automated data repositories. Consideration should be given to the placement of the data repository function with IACs because they have the trained staff to both handle numerical data and assess their worth.

c. Determine the function of DTIC with regards to use by IACs, as well as the DoD-wide user community. The ad hoc group observed that the IAC activities are still off to the side within DTIC. Further integration and emphasis of this essential information analysis function is important. In addition, the utility of DTIC towards general non-DoD users and IACs must be kept in mind and maximized to the extent possible within the same budget. For example, consideration should be given to the possibility of automatic distribution to IACs by DTIC of documents that are indexed as relevant. Currently, IACs frequently encounter difficulties in timely acquisition of hard copies on a specific request basis.

d. Determine the function of the IACs within the DTIC information mechanism; the quality of each individual IAC's output, the effectiveness of their dissemination methods, and efficiency of their operations. What is their value in terms of the total desired result? What is their value relative to technical evaluation effort and budgetary input (i.e. how do they rate in terms of cost effectiveness?).

e. Determine a priority ranking for data that are worth keeping and those that warrant further evaluation and dissemination in view of DoD technical needs as well as financial feasibility.

f. Identify to the extent possible DoD R&D research trends. This identifies trends for a future data user community for whom plans must be initiated today to meet tomorrow's needs.

g. Determine cost/benefit of data tagging and flagging for the DTIC bibliographic files. Would this increase DTIC usage and retrieval efficiency? To what extent should DTIC become involved at this stage? What would be alternate changes that could be made to assure retrieval of references to needed data?

The overall role of DTIC in numerical data must yet receive far more attention.

2. Quality of Data

Certain types of data require careful evaluation. For others, validation, appraisal, quality assurance are appropriate. A bibliographic information transfer system can at best regurgitate data of original reporting quality. In a case such as DTIC where most information is of report quality (i.e. not peer-reviewed) this means quite variable quality of any numerical data retrieved from the system. To remedy this situation, steps can be taken at several levels of sophistication.

a. Review DoD needs for evaluated data, for literature reviews, or data banks, e.g. through studying scientist's use of review literature as well as evaluated versus unevaluated data before initiating own laboratory work.

b. Review Information Analysis Centers. One effect of IACs whose products include even a small amount of mediocre output is that they tend to create a negative attitude from potential users towards all IAC outputs. Such attitude has the net effect of making accessible, evaluated data inaccessible. This situation must be avoided. It must be realized in this context that good-quality data compilations

require firm and uninterrupted funding: technical experts with special analytical skills are required for the task, but are difficult to recruit. Threats of sudden withdrawal of funding can do irreparable damage. If funding is to be reduced on a permanent basis, careful considerations are prerequisite, including user community input, in determining program reduction or redirection.

c. Review DoD laboratories and DoD contractors regarding individual organizational data quality assurance programs (do the laboratories have any in place?). Whenever possible, institutional quality assurance efforts should furthermore be closely linked with the laboratory research programs that generate and use the data.

d. Consider the possibility to increase the involvement of universities in the operation of the IACs either as direct contractors or as subcontractors. This would have the advantages of tapping a pool of trained talent, providing support for graduate students while training them in data quality matters (whether applied as future evaluators or as laboratory researchers), and adding the academic community to the list of supporters.

e. Develop criteria that can distinguish between the different levels of quality assessment or evaluation for the presentation of different types of data. An example is the data levels assigned in geodata to the different levels of data digestion and condensation. This relates to archiving of significant data: what should be retained, what discarded; unique events versus controlled, reproducible data. Who should have responsibility for archiving and accession and how should this be done? (This has an obvious relationship to the proposed data repository system as well as to the items above.)*

---

* The NRC's National Materials Advisory Board, Committee on Materials Information for Structural Design Used in Computerized Analysis and Manufacturing, is currently addressing the development of criteria for the presentation of data for engineering alloys. In this area, a tremendous amount of data is required to characterize a material for use in critical structures (a fact not commonly understood by those who control the funding of such information analysis activities). A first step in the development of criteria in this field would involve development of uniformity of data format. This study is being done under DoD and NASA sponsorship.

f. In development of an indexed directory of databases, the extent of evaluation or quality of data can be indexed. This requires some degree of indexer training.

g. Develop an indexing system (with or without data tagging and flagging) capable of notation as to whether data are reported together with discussion of instrumentation, errors, or ambient information.

3. Some specific data issues.

There are a number of data issues that are of importance at this time, and are elements of one or more of the above topics.

a. Determine the implications of the copyright law, transborder data flow and privacy of data to DoD information needs. Several ramifications of the new copyright law with regards to data flow are not yet clearly visible. A number of private organizations are attempting to gain insight in this matter. While NDAB does not now have a member expert, it is in contact with this community and could assist DTIC in establishing discussion and understanding in copyright questions.

b. Determine how to protect against the loss of data either by mechanical catastrophies or institutional changes and develop methods for reducing unwanted losses. Specifically, deliberate decisions must be made in 1) necessary redundancies, and 2) procedures in archiving. Expenses in carrying out data protection procedures and advantages of improved technologies are elements in decision making. Causes would include: data centers or information systems ceasing to exist; inadvertent tape erasures, fires, unreported results of failed experiments, changes in data coding and entry, and inadequacy of data reporting and documentation (it has been reported that the latter can cause losses for certain subjects of between 50 and 80% of the data reported in the literature).

4. Effective use of computer methods.

DTIC must maintain state-of-the-art methods due to the huge volumes of data processed, diverse user community, and need for prompt, accurate systems response. Topics for which answers must be found include:

a. What are best options for hardware and software for automation of databases?

b. How to achieve compatibility of databases and interfacing for interdisciplinary and interservice use? Compatibility factors include choice of characters, choice of system, method of data accumulation and analysis, method of expression of results, uniformity of standards and units used, standard data nomenclature and data elements, criteria for entries concerning quality, usage, needs and dissemination methods.*

c. To what extent can technology be transferred from existing information systems to those being created in new fields?

d. How humanly user effective are machine data systems for decision making in the military-scientific-industrial complex?

---

* The previously mentioned Committee of the National Materials Advisory Board expects to describe methods to achieve database compatibility for engineering alloy data in its final report.

Summary

In summary, this report identifies a number of the most significant points for DTIC consideration. The report is a preliminary one, based on a concise overview given by DTIC staff. NDAB activities such as reviews of DoD's present and future data requirements, reviews of quality of Information Analysis Centers, or specialized studies of data/information/intelligence capabilities for the DoD information microcosm, are at this point beyond the means but quite within the interest of the Board. NDAB welcomes discussion of any of the topics this report addresses at its future meetings, or with individual members, as time permits. However, more investigative, special studies would require an effort on NDAB's part outside its current level of activity.

It is hoped that DTIC, as the central DoD service unit for scientific and technological information, will actively enhance its numerical data capabilities to take a lead in DoD--as well as several national data handling techniques. NDAB has observed in the past that the numerical data quality aspect of this function was not given high enough importance and visibility within DoD. To enhance its program's effectiveness, it must receive such internal administrative recognition. (1)

Acknowledgement: This report has benefited from interaction of the NDAB work group with the NMAB Committee on Materials Information for Structural Design Used in Computerized Analysis and Manufacturing. While that study addresses one technical area only, several aspects necessarily interrelate closely with the broader issues addressed here by NDAB.

---

(1) This point was also addressed in the "Technical Information Conference for R&D Managers--Information Issues for the 80's", 16-17 March, 1981, National Defense University, Fort McNair, Washington, D.C.